Adobe

NASA/FAA Workshop V
July 24, 2002

Going Digital: Why Use Industry Metadata Standards … and How

Building an eXtensible Metadata Platform

Andrew Salop
Sr. Product Manager
Content Asset Management

www.adobe.com/products/xmp

The Landscape: Network Publishing

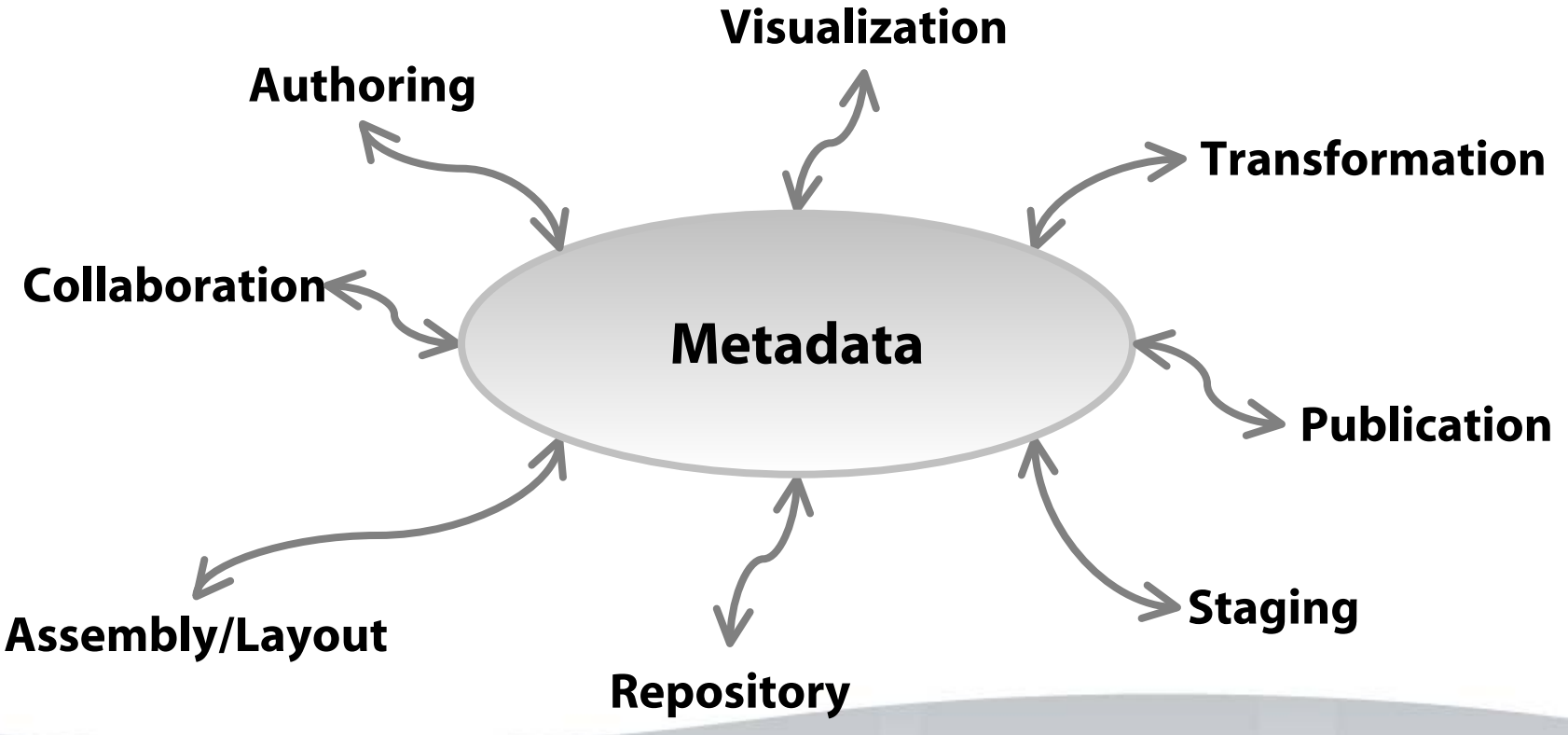
- **The third revolution in publishing technology, after …**
  - Desktop Publishing -1980s (digital production)
  - Web Publishing -1990s (mass-media digital publishing)

- **Making compelling content reliably available any time, anywhere, on any device**
  - Re-use, Re-purpose, Re-express
  - Print, Web, ePaper, handheld, broadcast, eBook, etc.)
  - Text, image, graphic, 3D, video, audio, animation, template
  - Files, objects, assets, documents, content, "data"
    - Individual and compound

- **Create, Manage, Deliver**

Network Publishing Ecosystem

Metadata interchange is fundamental to content management & Network Publishing



Metadata Describes Intent

- Data about data
- Not content

- Analogies
  - Label on a can
  - Annotations on pictures
  - Name tag

Metadata Categories

Implicit(generated) vs. Explicit(entered)

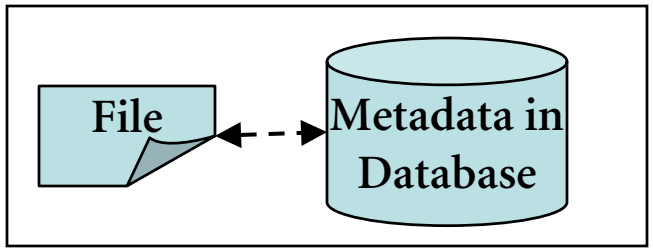
- **Identification**
  - Unique characteristics facilitating lookup and retrieval
  - title, creator, publisher, format
- **Discovery**
  - Searching, Querying, Categorization
  - subject, description, keywords
- **Rights**
  - Ownership, permission, access (copywrite)
- **Relational**
  - Describes relationship with other elements (rendition of)
- **Temporal**
  - Parameters to judge relevance
  - creationTime, publicationTime, expirationTime
- **Structural**
  - Describes logical structure of content (headline, byline)

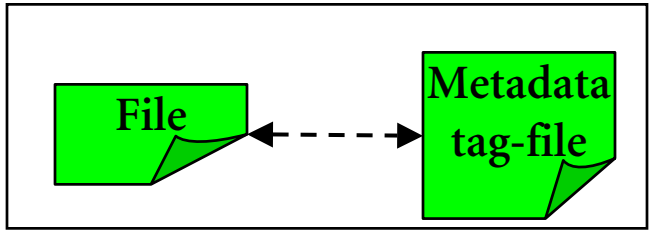
Metadata Dimensions

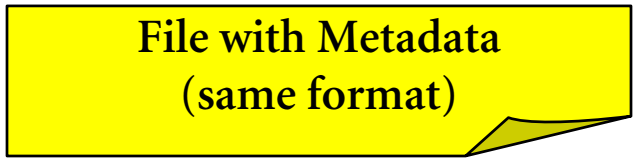
- **Schema**
  - Properties, vocabularies (controlled or otherwise)
  - Often provided through metadata standard bodies that address a particular domain
  - (defined by humans)
- **Encoding / Framework**
  - How metadata is expressed (XML, binary, etc.)
  - Usually through a defined format or in a database
  - How read and written by devices (SW&HW)
- **Logistics**

Metadata Logistics

- **Metadata in database with semantic link to separate file**
  - May be separated
  - "smart database, dumb file"
- **Separate Tag File, linked to content**
  - Can also be separated
- **Incorporated into file format**
  - File format may be difficult to read outside of authoring application
- **Embedded metadata format, within file format**
  - Useful if metadata is encoded in open standard format (such as XML)
  - "smart" file/asset/object/document







Metadata "Standard" Segmentation

Domain specific

- Library Science / Discovery
- Digital Photography
- News syndication
- Learning
- Video/broadcast
- eBook
- Magazine
- Enterprise / Proprietary
- … others

No common encoding mechanism

Metadata Workflow Issues

- **Capture**
  - Accuracy, consistency
- **Compatibility**
  - Dropped or locked metadata
- **Overlapping vocabularies (schemas)**
  - Colliding properties (e.g., author, creator)
- **Reconciliation**
  - Updating to reflect a change in a file's status or characteristics
- **Interoperability**
  - Preservation & propagation across multiple file formats and databases
  - Translating industry specific terminology

Why Use Standards?

- **Schema**
  - Rather than building from scratch …
  - leverage vocabularies that reflect prior thought,
  - then tweak to meet your needs
- **Encoding / Framework**
  - Think Network Publishing, i.e., multiple domains
  - Use an encoding mechanism that preserves and propagates metadata across multiple domains
- **Logistics**
  - Maintain linkage between data & metadata
  - Keep your assets “smart.”

Customer Requirements:

Open	Build on standards
Extensible	Support custom vocabularies
Universal	Support diverse file formats: JPEG, GIF, TIFF, EPS, PDF, HTML, XML, PNG, PSD …
International	Unicode, multilingual
Robust	Keep metadata with asset

Adobe

XMP

Award winning

file format neutral

extensible

metadata platform

Adobe and Metadata

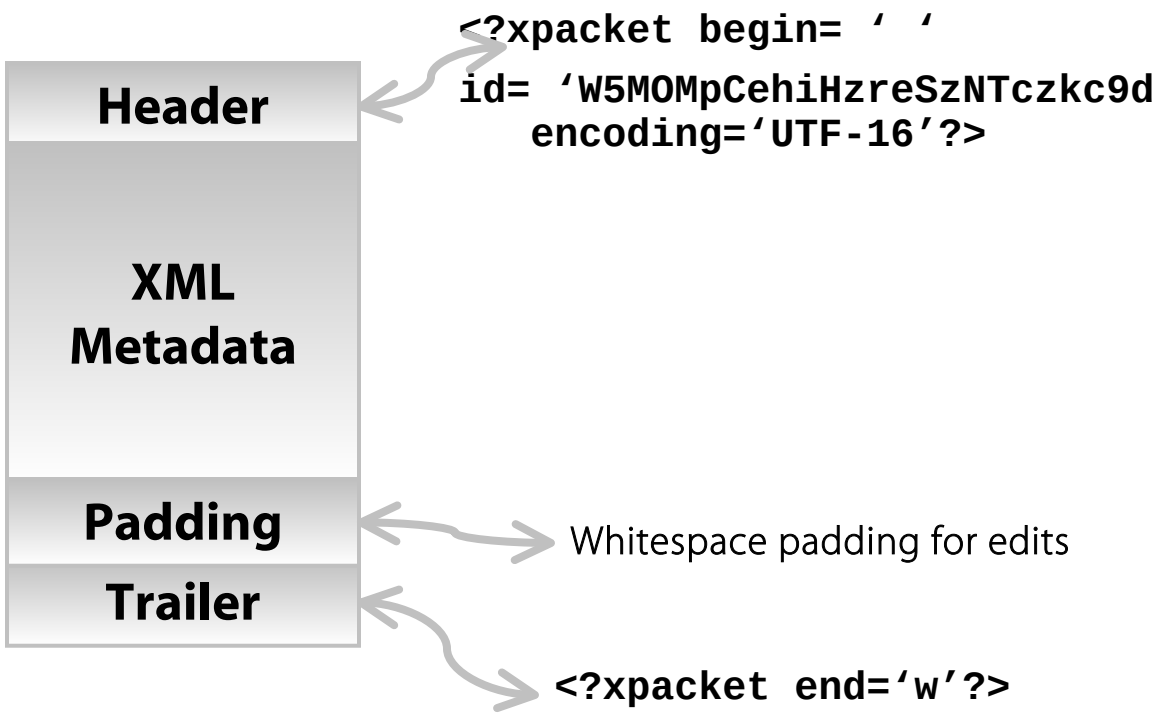
- **XMP is Adobe's preferred metadata format**
  - Because it is extensible
    - Schema neutral (important for asset management)
    - File format neutral (important for workflows)
- **Integrating into Adobe software applications**
  - Common user experience across diverse tool & media types
    - Imaging, graphics, layout, ePaper, etc.
  - Easily supported by devices, software, formats, institutions
- **Co-exist with existing workflows**
  - Support for legacy standards
  - Common encoding specification

XMP Components

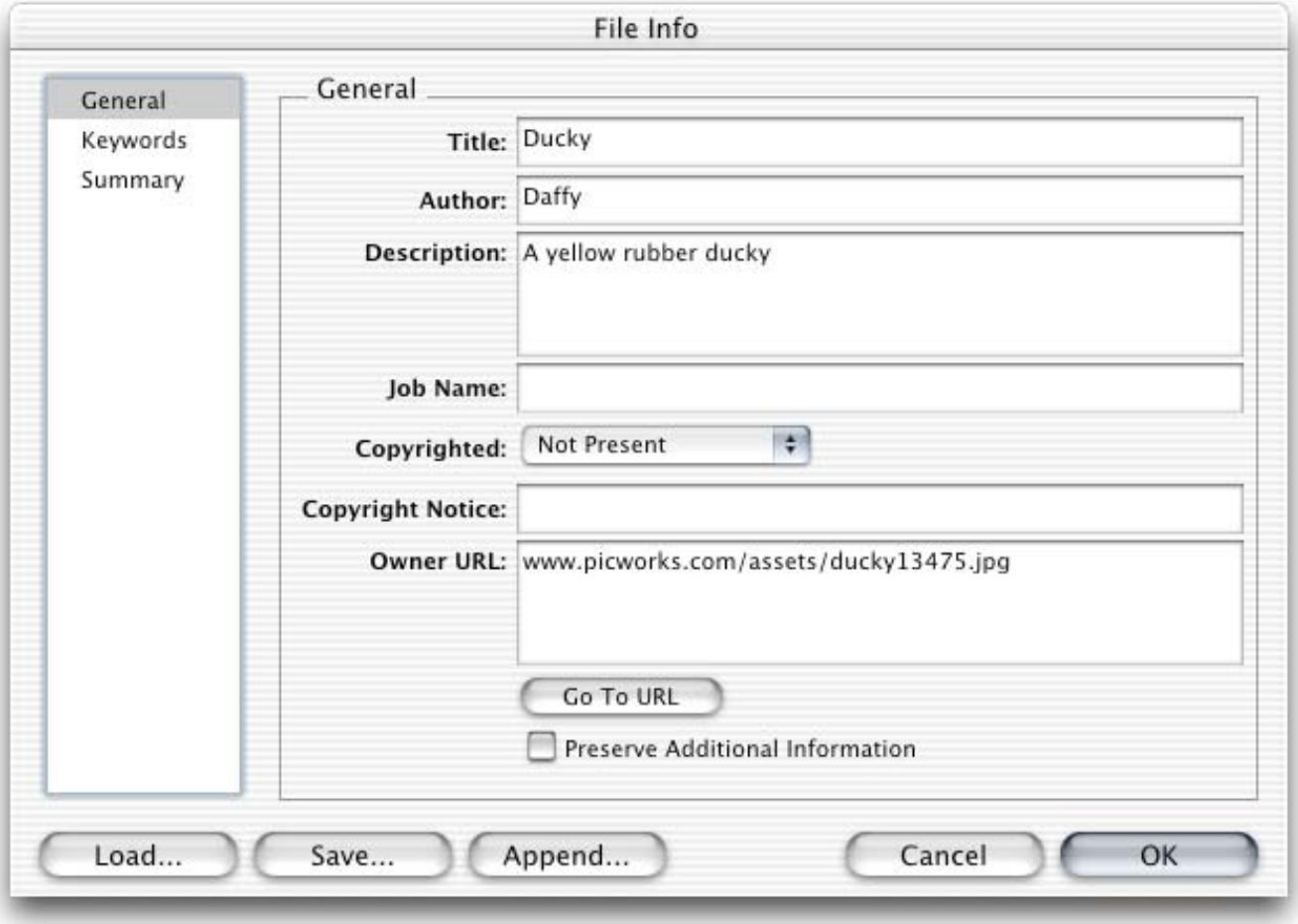
1. **XML Packet**
   - Embeds XML within binary objects
   - Accessible outside authoring application
2. **Framework**
   - Based on W3C standard for XML Metadata (RDF)
   - Supports aliasing of equivalent properties
   - Works within and external to file formats
3. **Schema (list of properties)**
   - Core (based on Dublin core)
   - Extensible
4. **Software Developer Kit (SDK)**
   - Released September 2001
   - Free download: (www.adobe.com/products/xmp)

XML Packet

The secret to embedding XML in any file



Common implementation across Adobe applications



XMP in Adobe Products

Acrobat

GoLive

Illustrator

InDesign

LiveMotion

InCopy

FrameMaker

Photoshop

Altercast (next)

Industry Support

- Artesia
- Documentum
- Getty Images
- IBM
- Interwoven
- Kodak
- KPMG Consulting
- MediaBin
- North Plains Systems
- W3C
- WebWare
- Xerox

Benefits

- **Enables creation of "smart assets"**
  - promoting re-use, re-purposing & re-expression
- **Standardized tagging framework**
  - promotes metadata capture, preservation and propagation
  - across devices, software applications, file formats, institutions
  - works across network publishing workflows / create-manage-deliver
- **Extends metadata beyond the context of a database**
- **Embeds access to rights information**
  - Provides both absolute and relative references
- **XML metadata accessible without understanding host file format**
- **Open specification, based on industry standards**

XMP Resources

www.adobe.com/products/xmp

- **Data sheet**
- **White Paper**
- **Online knowledgebase**
- **Software Development Kit (SDK)**
  - Available via open source license
  - Contributors free to submit requests, bug fixes, suggest changes

Adobe

SEYBOLDREPORTS.com

SEYBOLD EDITORS' HOT PICKS San Francisco 2001

"Of all the announcements, the one that's least visible [XMP] is likely to have the biggest long-term impact …

… Adobe is the first vendor to announce a framework for embedding metadata across multiple desktop applications, and XMP is sufficiently open and nonproprietary that it deserves consideration as an important new industry standard."

The Bulletin: Seybold News & Views On Electronic Publishing